University of Gondar

College of medicine and health science

Institute of public health

Predictors of CD4 count change over time among HIV positive children receiving Antiretroviral Therapy in Amhara region, 2017. Retrospective longitudinal data analysis

BY: Tilahun Yemanu (BSc)

Name of advisors:

1. **Lemma Derseh (MPH, Asst. professor)**
2. **Destaw Fetene (BSc, MPH)**

ATHESIS SUBMITTED TO THE INSTITUTE OF PUBLIC HEALTH, COLLEGE OF MEDICINE AND HEATLH SCIENCES, UNIVERSITY OF GONDAR IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTTER OF PUBLIC HEALTH IN BIOSTATISTICS.

Gondar, Ethiopia

June 2017

University of Gondar

College of Medicine and Health Science

Institute of public Health

Department of Epidemiology and Biostatistics

Predictors of CD4 count change over time among HIV positive children receiving Antiretroviral Therapy in Amhara region, 2017. Retrospective longitudinal data analysis

By: Tilahun Yemanu (BSc)

Email: yemanu.tilahun@gmail.com

Phone: 09 25 97 84 89

Approved by the Examining Board

_____ _____

Head, Department of Epidemiology and Biostatistics

Advisors:

1._____ _____

2._____ _____

_____ _____

_____ _____

Examiners

Acknowledgement

I would like to thank University of Gondar College of Medicine and Health Sciences for their financial support and department of Epidemiology and Biostatistics for opening this masters program in order to apply application based education in real data set in health problems.

Next, I would like to express my sincere gratitude to my advisors Mr. Lemma Derseh and Mr. Destaw Fetene for their continuous comments, motivation, and support. The guidance of them was helping me throughout the time to accomplish this work.

Lastly but not the least I would like to thank my classmates helping me from initiation up to finalization of this work

List of acronyms/Abbreviations

3TC	Lamivudine
ABC	Abacavir
AIDS	Acquired Immune Deficiency Syndrome
AIC	Akaki Information Criteria
ART	Antiretroviral Therapy
ARV	Anti-Retroviral
AZT	Zidovudine
BIC	Bayesian Information Criteria
CSA	Central Statistical Agency
d4T	Stavudine
EFV	Efavirenz
HAART	Highly Active Antiretroviral Therapy.
HIV	Human Immunodeficiency Virus
LPV/rtv	Lopinavir/ritonavir
LMM	Linear Mixed Model
ML	Maximum Likelihood
OIs	Opportunistic infections
REML	Restricted Maximum Likelihood
TB	Tuberculosis
TDF	Tenofovir
UNAIDS	United Nation Program on HIV/AIDS
VL	Viral load
WHO	World Health Organization

Table of Contents page

Acknowledgement ii
List of acronyms/Abbreviations iii
List of table vi
List of figure vi
Abstract vii
1. Introduction 1
1.1 statement of the Problem 1
1.2 Literature review 3
1.2.1 Evolution of CD4 cell count 3
1.2.2. Determinants of CD4 count change 3
1.2.2.1. Socio demographic and clinical factors 3
1.3 Justification 6
2. Objective 7
2.1 General objective 7
2.2 Specific objectives 7
3. Methods 8
3.1 Study area 8
3.2 Study design 8
3.3 Source population and study population 8
3.4 inclusion and exclusion criteria 8
3.4.1 Inclusion criteria: 8
3.5 Sample size determination 9
3.6 Sampling procedure 9
3.7 Study variables 11
3.7.1 Dependent variable 11
3.7.2 Independent variable 11
3.8 Operational definitions 11
3.9 Data collection instruments & procedures 11
3.10 Data quality assurance 11
3.10 Data processing & analysis 11
3.12 Statistical analysis 12

3.12.1 Exploratory data analysis 12
3.12.2 Linear mixed models 12
3.12.2.1 Univariate Linear Mixed Model 13
3.12.2.2 Multivariate Linear Mixed Model 13
3.13 Model Diagnostics 13
4. ETHICAL CONSIDERATION 14
5. Results 15
5.1 Sociodemographic, medication related, clinical and behavioral factors 15
5.2 Exploratory data analysis 17
5.2.1 Individual profile plot for some selected patients 17
5.2. 2. Exploring the mean structure 18
5.2.3 Exploring correlation structure 21
5.2 Linear Mixed Model 22
5.2.1 Univariate Linear Mixed Models 22
5.2.2 Multivariate Linear Mixed Model 23
5.3 Model diagnostic checking 26
6. Discussion 29
7. Conclusion and recommendations 32
7.1 Conclusion 32
7.2 Recommendations 32
Reference 33
Annexes I 36
Annexes II. Information Sheet and Consent Form 38
Annexes III: Stata codes 40
Annexes IV: some selected stata outputs 42

List of table

Table 1: Sociodemographic, medication related, clinical and behavioral characteristics ART data set taken in Amhara region Hospital from 2010 –2016 15
Table 2: Descriptive Statistics of Continuous Covariates at baseline for ART data set taken in Amhara region from 2010-2016 16
Table 3: Mean, standard deviation, and Number subjects in each time for ART data set taken from Amhara region from 2010-1016..... 19
Table 4: *selection of correlation structure* for ART data set taken from Amhara region from 2010-2016 22
Table 5: Random Effects Models with the associated values for the likelihood ratio test and p-value for ART data set taken from Amhara region from 2010-2016 23
Table 6: tests for model extension for ART data set taken from Amhara region from 2010-2016 23
Table 7: *multivariate analysis of linear mixed model for ART data set taken from Amhara region from 2010-2016* 25
Table 8: Random parameter estimates for ART data set collected in Amhara region from2010-2016. 26

List of figure

Figure 1: conceptual frame work of CD4 count change among HIV infected children receiving ART in Amhara region 5
Figure 2: diagrammatic representation of sampling procedure 10
Figure 3: Individual profile plot for logCD4 count for ART data set taken from Amhara region from 2010-2016 17
Figure 4: Individual profile plots for logCD4 count grouped by gender for ART data set taken from Amhara region from 2010-2016..... 18
Figure 5: Smoothing Mean profile plots for the logCD4 count for ART data set taken from Amhara region from 2010-2016. 20
Figure 6: Smoothing mean profile grouped by gender for ART data set taken from Amhara region from 2010-2016 20
Figure 7: Scatter plot of correlation matrix for ART data set taken in Amhara region from 2010-2016 22
Figure 8: residual plot for logCD4 count for ART data set taken from Amhara region from 2010-2016 27
Figure 9: Smoothing mean profile by ART adherence for ART data set taken from Amhara region from 2010-2016 43

Abstract

Background: Chronic HIV infection results in the progressive depletion of CD4+ T lymphocytes from both lymphoid tissues and peripheral blood. The existing studies tried to address the evolution of CD4 cell count by using logistic and linear regression. However, these models does not show the change of CD4 cell count over time rather the effect of predictors on the evolution of CD4 cell count on the given point of time and it gives biased inference for the estimates. This study designed to addresses this problem by using longitudinal data analysis which accounts both fixed and random effects on the model.

Objective: The aim of this study was to assess the CD4 count change over time and its predictors among HIV positive children receiving antiretroviral therapy in Amhara region, Ethiopia.

Methods: A retrospective follow up study was conducted among children who received Antiretroviral Therapy between 2010 and 2016 in Amhara region public Hospitals. Systematic random sampling was applied to select the 936 patients chart from the ART registration book. The data were extracted from the selected chart by the trained data collectors. The data were entered in to Epi info 7 and exported to stata14 for further analysis. Since repeated measures were taken, correlation was taken in to account when analyzing the data. Due to this illustration Linear Mixed Model was applied to describe CD4 count change over time for children on ART.

Result: A total of 936 HIV positive patients were followed retrospectively with mean CD4 count at baseline of 465.1cells/mm^3 and an average rate of CD4 count change 5.01cells/mm^3 per month. A patient having opportunistic infection (β = -0.048, 95% CI -0.092, -0.0035), disclosed their status (β= -0.088, 95% CI -0.135, -0.041) and baseline WHO clinical stage II (β = -0.0898, 95% CI -0.134, -0.046) are determinant predictors of CD4 count change.

Conclusion: the average rate of CD4 count change over time was 5.01cells/mm^3 per month**.** The CD4 count change over time was described by quadratic time function. In addition baseline WHO clinical stage, opportunistic infection, disclosure and regimen type was significant predictors for CD4 count change.

Key words: CD4 count, Linear Mixed Model, Antiretroviral Therapy

1. Introduction

1.1 statement of the Problem

HIV/AIDS is one of destructive diseases human kind ever faced. It brings with it profound social, economic, and public health consequences. It has become the world's most serious health and development challenges. HIV is a leading cause of death worldwide. In addition Paediatric HIV has gone from being universally fatal to becoming a treatable chronic condition[1, 2].

Among infants and children not taking ART, HIV infection is often rapidly progressive and fatal. Approximately 20% of HIV-infected infants will die by 3 months of age without treatment, half will die before reaching their second birthday, and three-fourths will die by 5 years of age[3, 4]. Among those infected children 90% are found in sub Saharan Africa[5].

According to Ethiopian public health institute ministry of health report on HIV related estimate and projection in 2014 there are total of 95,795 infected children in Ethiopia in 2016 from this 48,411 are male and 47,384 are females. In addition from this year in Ethiopia there are a total of 1,408 new infection of children and 1,459 annual related deaths of children in this year[6].

Chronic HIV infection results in the progressive depletion of CD4+ T lymphocytes from both lymphoid tissues and peripheral blood. Thus, the monitoring of peripheral blood CD4 cell count is the standard used in decision-making concerning initiation of antiretroviral therapy (ART), as well as monitoring response to ART over time[5, 7, 8].

The children with low CD4 count at pre-ART level initiated with HAART are more prone to opportunistic infections, AIDS defining illness like malignancies, low BMI, decreasing body weight, advanced clinical stage, poor phenotypic appearance, growth retardation and lipoatrophy and dystrophy[9].

Combination of antiretroviral therapy (ART) is the only effective treatment available for the suppression of HIV, and early ART initiation significantly reduces AIDS-related initiation significantly reduces AIDS related morbidity and mortality among children[10].

CD4 cell count and HIV RNA viral load in response to antiretroviral treatment are important measures of the efficacy of ART in individual patients and the effectiveness of ART in population of pediatric patients enrolled in HIV care and treatment programs[11]. Studies conducted on association between combination of ART and evolution of CD4 count on children [12-14]. However, the majorities of existing studies are based on small number of participants and short duration of follow up as well as single centered study. are does not show the improvements of CD4 cell count over time it simply shows the effect of predictors on the evolution of CD4 cell count on the given point of time because simple regression and logistic regression models for repeated data are inadequate for model fitting and can lead to inappropriate conclusions and inference.

In addition, there is limited information regarding the application of statistical models to predict AIDS disease progression using longitudinal CD4 cell counts in Ethiopia.

The increasing number of HIV-infected children starting antiretroviral therapy and the limited capacity to model CD4 cell count using longitudinal data analysis is the main interest to indentify evolution of CD4 cell count in this study. The application of modeling technique was essential due to repeated nature of data.

1.2 Literature review

1.2.1 Evolution of CD4 cell count

Study conducted in Bangalore and west Africa showed that early inception of HAART mainly helps the children to maintain better CD4 count with good immune systems and they are less prone to OI's and it also reduces the progression of disease, mortality and RNA plasma viral load[9, 15].

Study conducted in Zambia [16] on the objective of to report early clinical and immunologic outcomes of children enrolled in the pediatric treatment program discuses that children receiving ART showed significant improvement in CD4 cell count, weight gain, and hemoglobin concentration.

Studies done on predicting patterns of long-term CD4 reconstitution in HIV-infected children starting antiretroviral therapy in sub-Saharan Africa[17, 18] was briefly discussed that the majority of children (76%) had asymptotic CD4 reconstitution following ART initiation, with an initial steep increase in CD4 cell count for age that slowed with half life of 18 weeks tending towards a constant level over a long term of (for a child of average demographics) , 80% of the CD4 count expected in an HIV-uninfected child of the same age. CD4 for age both before ART and in the long term were higher for children starting ART at younger ages, and long term CD4 reconstitution was better in children with higher CD4 for age at ART initiation.

1.2.2. Determinants of CD4 count change

1.2.2.1. Socio demographic and clinical factors

study conducted on the objective to assess whether treatment of HIV infected children with HAART would improve their immune function[13, 19] result showed that there is a significant increase in the peripheral CD4 T-cell count at week 24 (median change, 231 cell/µl; range -33 to 548/µl [p=0.0029,wilcoxon signed rank test]) was accompanied by a substantial decrease in the HIV load (median change, -1.64$\log_{10}$ copies/ml; range -2.98 to -0.12 $\log_{10}$ copies/ml [p = 0.001]).

Study conducted on food insecurity and CD4% among HIV positive children in Botswana[20] showed that food insecurity was inversely associated with CD4%,

adjusting for covariates, among HIV positive pediatric patients in a sub-Saharan African setting, the region with the highest prevalence of HIV in the world.

Studies conducted in Zimbabwe and Democratic republic of Congo[21, 22] showed that the median CD4 cell count at different time points shows an increasing trend in the levels of CD4 cell count with time for both age groups. From those age group and baseline CD4 count were significantly associated with an increase in the levels of CD4 cell count over time.

A study conducted on Absolute CD4+ T-Lymphocyte Count as a Surrogate Marker of Pediatric HIV Disease Progression[8] the result shows that there was a significant effect of treatment regardless of the regimen on CD4+ T-lymphocyte counts ($p < 0.001$) and participants who had other infections or opportunistic infections had lower levels of CD4+ T-lymphocyte counts.

A longitudinal single center observational study [23] supports that HAART induces a beneficial effect in terms of clinical, virologic and immunologic outcomes even in children previously exposed to antiretroviral therapy. They also found that CD4 T cell increased significantly in most children (65%) irrespective of the extent of virologic suppression and baseline age of children, baseline CD4 count and the number of ARVs was found to predict immunological response.

A study conducted in India [24] discuses that non adherence children was associated with loss of CD4 count and increase the viral suppression of immune system, while a an adherent children was associated with CD4 gain, reduces drug resistance and viral load suppression.

Studies conducted in Brazil on choosing the right strategy to model longitudinal count data in epidemiology with application of CD4 cell counts [25] discuses that analysis of longitudinal data using conventional regression models is inadequate as they fail to consider the dependence between observations over time. Longitudinal data may also present additional complexities in its structure, which may occur due to the imbalance and/or the fact that they are unevenly spaced, or owing to missed data. It is up to the date analyst to conduct a thorough exploratory analysis to evaluate the data structure and choose the statistical model that best suits it.

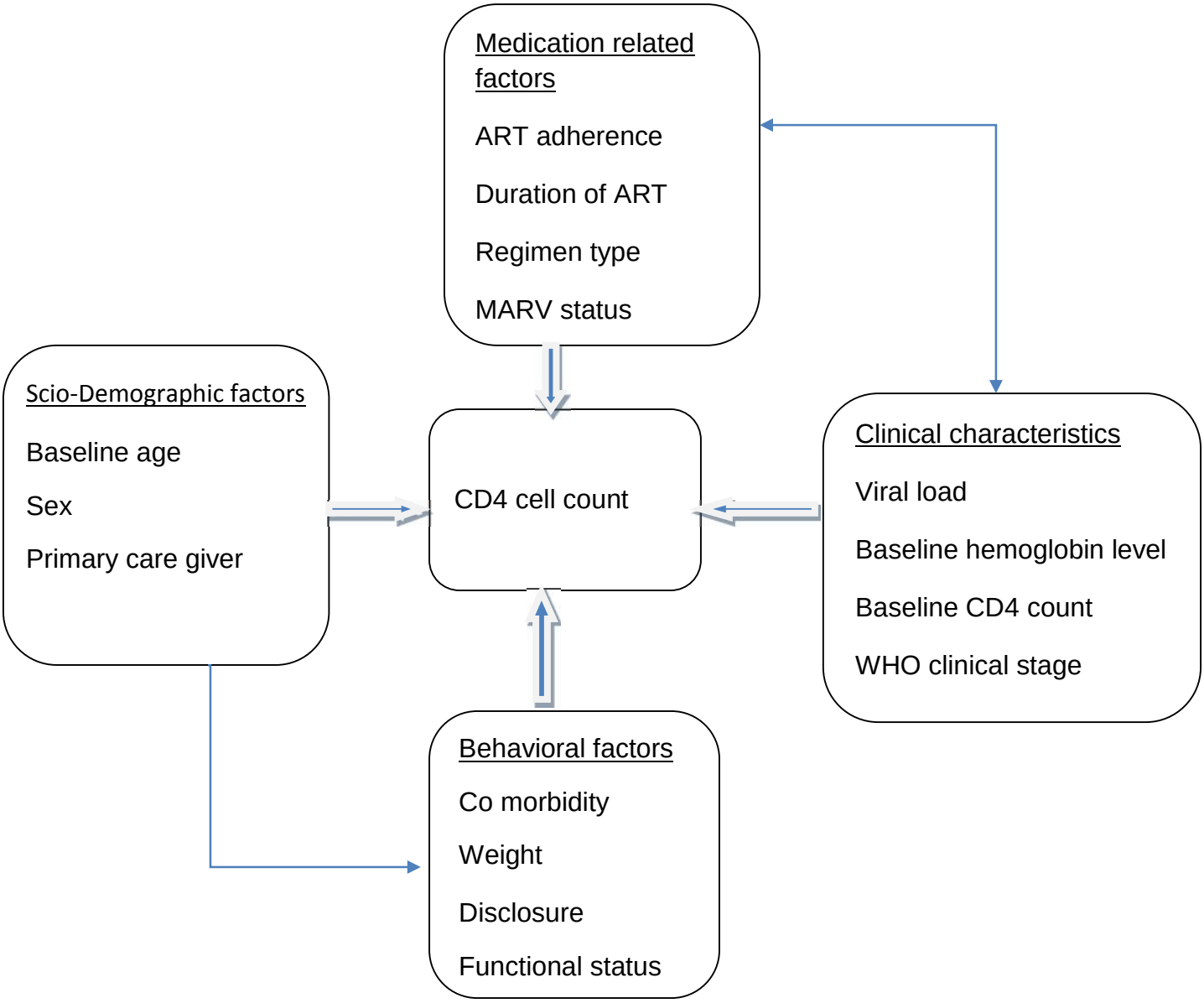


Figure 1: conceptual frame work of CD4 count change among HIV infected children receiving ART in Amhara region

1.3 Justification

The CD4 cell count is a critical measure of immune system in HIV patients and is used as an imperative biomarker for describing the progress to AIDS in children's and adults. Studies were conducted on the effects of ART on evolutions of CD4 cell count among HIV positive children. However, the majorities of existing studies were without considering the evolution of CD4 cell count over time and based on the assumption that data are independent over time and this leads to the use of logistic and linear regression model, but those models are inadequate for fitting correlated data and this could leads to biased statistical inference. Hence, further study on the evolution of CD4 cell count using linear mixed models are needed for longitudinal data analysis. This study is designed to examine the CD4 count change over time and to identify its predictors among HIV positive children on ART in public Hospitals of Amhara region, Ethiopia.

.

2. Objective

2.1 General objective

The aim of this study was to assess the CD4 count change over time and its predictors among HIV positive children on antiretroviral therapy in Amhara region public Hospitals.

2.2 Specific objectives

- To examine the CD4 count change over time among HIV positive children on Antiretroviral Therapy.
- To identify predictors which determine the CD4 count change among HIV positive children on Antiretroviral Therapy.

3. Methods

3.1 Study area

The study was conducted in Amhara regional state public Hospitals. The region is found in Northwest Ethiopia which is the second most populous with 19,602,512 estimated total populations according to CSA 2011 report. The region have 19 public hospitals, 801 health center, 3302 health posts, and 1031 private health facilities (clinics and hospitals). All the public health Hospitals provide ART service for both adults and children of the people of Amhara and neighboring regions. According to regional health office report over 2,112 pediatric patients are actively enrolled on ART in those hospitals and clinics currently. From these about 70% of all ART services covered by the six hospitals namely; Gondar university teaching and referral hospital, Debre markos, Debre Birhan referral, Dessie, Debre tabor and Felegehiot specialized hospitals.

3.2 Study design

A retrospective follow up study was conducted during 2010- 2016 from ART clinic in Amhara region.

3.3 Source population and study population

The source populations for this study were all HIV positive children from under the age 15 years that follow first line ART.

The study populations were all children that follow first line ART whose age is less than 15 years in Amhara region.

3.4 inclusion and exclusion criteria

3.4.1 Inclusion criteria:

All HIV positive children who started first line ART treatment and have at least two CD4 measurements were included in this study.

3.5 Sample size determination

The sample size for this study was done based on the following assumption

All subjects measured at m = 8 time points (no drop-out)

Constant within-subject correlation ρ = 0.5

σ^2 =1

We want 90% power to detect a difference d of 0.25 at the two-sided 0.05 significance level.

$$n = \frac{4\sigma^2(1+(m-1)\rho)(Z_{\frac{\alpha}{2}}+Z_{1-\beta})2}{md^2}$$

$$n = \frac{4(1+(7)0.5)(1.96+1.645)2}{8(0.25^2)} = 467.86 \approx 468$$

Therefore, the normal sample size calculation was give 468, but by adjusting design effect 2 we get total sample of 936 subjects. Detailed sampling technique shown in the next figure.

3.6 Sampling procedure

The sample of this study was performed on the selected major hospitals based on their considerable coverage of patients served and independent probability sampling method was conducted in order to select patients chart from patients' registration book.

A proportional allocation was employed to those hospitals in order to obtain the best precision of the outcome. Systematic random sampling technique was used to select patient review chart from Gondar university teaching and referral, Felegehiot, Debre markos, Debre Birhan, Dessie, and Debre tabor hospitals.

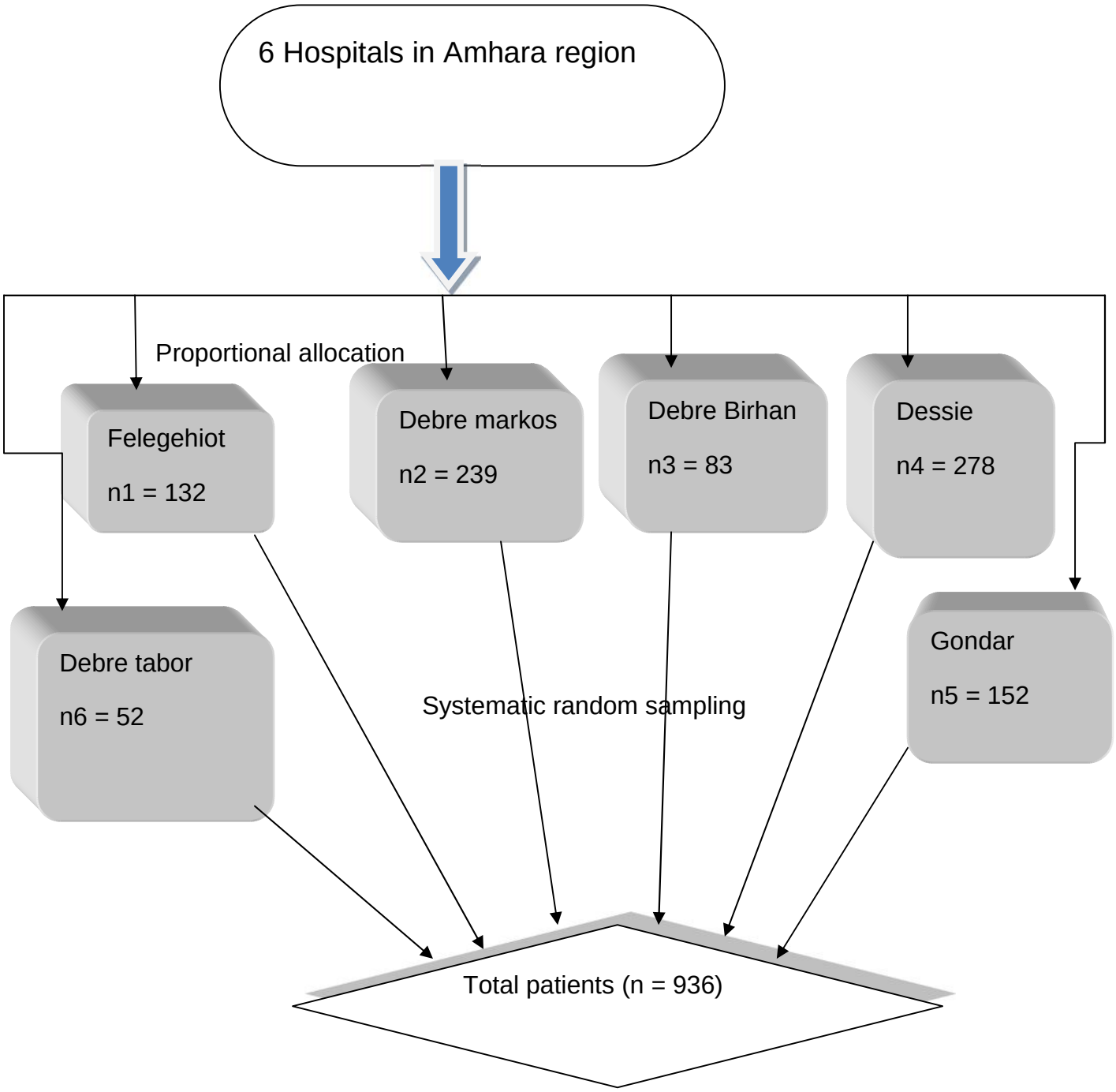


Figure 2: diagrammatic representation of sampling procedure

3.7 Study variables

3.7.1 Dependent variable

CD4 count

3.7.2 Independent variable

Explanatory variables are sex, age, baseline CD4 count, baseline weight, regimen type, WHO staging, functional status, baseline hemoglobin level, disclosure, mothers ARV status and time in months.

3.8 Operational definitions

Adherence: a child said to be adherent if she/he delayed from clinical attendance no more than 3 days (visit more than 95% of appointment correctly) for each month. Good adherence is >95%, fair adherence 85%-95% and poor Adherence <85% delayed from clinical attendance (appointment) over the duration of the follow up.

3.9 Data collection instruments & procedures

The data collectors for this study were BSc Nurse working in the ART Clinic using uniform data abstraction format prepared for this study.

3.10 Data quality assurance

The quality of data was ensured through training of data collectors and supervisors, close supervision and prompt feedback. The training consisted of instruction on extracting technique, a detailed review of the abstract form content with practical demonstration.

The data was checked for any inconsistencies, coding error, out of range, completeness accuracy, clarity and appropriate correction was made by principal investigator and supervisors consistently on the daily basis.

3.11 Data processing & analysis

The data was entered in to Epi Info 7 and exported to stata14 for further analysis.

CD4 cell count at follow up time was described using the mean and standard deviation. Frequencies were used to describe categorical variables like WHO clinical stage, gender, opportunistic infection and primary care giver.

3.12 Statistical analysis

3.12.1 Exploratory data analysis

Exploration were done using graphical technique such as individual profile plot which gives us an idea on the within and between subject variability and mean profile plot that suggests the initial plausible assumption on the mean structure of model further more variance function that will suggest initial assumption on the structure of random effects.

Average profile plots were constructed to describe the mean evolution of CD4 counts, overall and according to different subgroups. From such exploration, indications were obtained, about the functional form of the evolution and also whether the evolution depends on given covariates. Smoothing using the Loess method was applied because the measurements were not equally spaced across the different subjects. To get rid of the skewness in CD4 data logarithmic transformation was applied and the analysis was carried on with the transformed outcome.

To select the covariance structure which provides the best fit for the data and thus reducing the risk of model misspecification, the four structures (identity, exchangeable, independent and unstructured) was fitted and the one with the smallest model akaki information (AIC) were selected. To determine the factors associated with changes in CD4 cell count univariate analysis for each baseline factor was assessed and those found to be significant (p-value<0.2) were selected in order to fit multivariable analysis.

3.12.2 Linear mixed models

Repeated measures of CD4 cell count were obtained every six months for 6 years in HIV positive children on ART. The Linear Mixed Model is one of the statistical methods used to analyze repeated measurements in continuous longitudinal data in an easy, valid and flexible manner[26]. The linear mixed-effects model is defined as,

$$Y_i = X_i\beta + Z_i\, b_i + \mathcal{E}_i\,, \qquad b_i \sim N(0, D), \mathcal{E}_i \sim N(0, \sum_i)$$

$$\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n \text{ are independent}$$

In which β is a vector of population-average regression coefficients called fixed effects, and where b_i is a vector of subject-specific regression coefficients? b_i Describe how the evolution of the i^{th} subject deviates from the average evolution in the population.

The matrices X_i and Z_i are $n_i \times$p and $n_i \times$q matrices of known covariates. The residual components ε_i are assumed to be independent N $(0;\sum_i)$, where $\sum_i$ depends on i only through its dimension n_i

3.12.2.1 Univariate Linear Mixed Model

The main purpose of fitting univariate linear mixed model in this study was to select important covariates used for fitting multivariate linear mixed model. The model development process was carried out by selecting covariates that have the potential to be included in the multivariate linear mixed model. Variable selection was done using the manual backward selection criteria. It was carried out restricted maximum likelihood method estimation and by using the selected correlation structure.

3.12.2.2 Multivariate Linear Mixed Model

In fitting the linear mixed model, a series of correlation structures of the longitudinal data of CD4 counts on HIV patient were considered. From the possible covariance structures in fitting the model, the one with the smallest AIC and convergence of the model in REML and ML was considered. Restricted Maximum Likelihood (REML) was preferred to Maximum Likelihood (ML) testing here, because it reduces the well-known finite down ward bias in the estimation of the covariance. Those, unstructured correlation structured was selected for assessing the change of CD4 counts by linear mixed model based on the smallest AIC and BIC.

3.13 Model Diagnostics

In model diagnostics for longitudinal data analysis used for the residuals are frequently used to evaluate the validity of the model assumption. In normal linear models residuals are used to verify linearity of effects, normality, independence and homoscedasticity of the errors and the presence of outliers or influent observations. The residual is the difference between an observed quantity and its estimated or predicted value. Two types of residuals were used in this study. Those are marginal residuals r_m and the conditional residuals r_c.

The marginal residual are $r_{mi} = y_i - X_i\hat{\beta}$ and the conditional residual $r_{ci} = x_i\hat{\beta} - z_i\hat{y}_i$

4. ETHICAL CONSIDERATION

Ethical clearance was obtained from the University of Gondar College of Medicine and Health science Institute of Public Health Ethical review committee. Letter of permission to use the data were obtained from the authority of each hospital administration in HIV care center. The study was conducted on secondary data so there is no harm to study participants because there is no direct contact with the subjects and also patients' identification numbers were used in the data set instead of names in order to keep their confidentiality.

5. Results

5.1 Sociodemographic, medication related, clinical and behavioral factors

Of the participants, 509 (53.4 %) were male and the primary caregiver for the children in were predominantly the biological mother (parents) 731 (76.6%). In addition about 286(29.9%) patients were found in WHO stage I. 323(33.86%) of patients had working functional status. 720(75.7%) were disclosed to the disease status while the rest 231(24.3%) were not. Furthermore, 691(72.8%) have other opportunistic infections (see table1).

Table 1: Sociodemographic, medication related, clinical and behavioral characteristics ART data set taken in Amhara region Hospital from 2010 –2016

Variable	Category	Frequency	Percentage
Sex	Male	509	53.5%
	Female	442	46.5%
Primary care giver	Parents	731	76.6%
	Grand parents	93	9.8%
	Relatives	68	7.2%
	Others	58	6.1%
Regimen type	d4T+3TC+NVP	327	34.3%
	d4T+3TC+EFV	76	8.0%
	AZT+3TC+NVP	333	34.9%
	AZT+3TC+EFV	160	16.8%
	Others	55	5.8%
Functional status	Working	323	33.9%
	Ambulatory	462	48.4%
	Bedridden	166	17.5%
BWHO stage	Stage I	286	29.9%
	Stage II	309	32.4%
	Stage III	287	30.1%
	Stage IV	69	7.2%
Disclosure	No	231	24.3%
	Yes	720	75.7%
Opportunistic infection	No	260	27.25%
	Yes	691	72.8%
MARV status	No	42	4.4%
	Yes	909	95.5%

The mean baseline age of children was 6.5 years with a standard deviation of 3.8 years; the mean baseline CD4 count was 465.1cells/mm^3 with a standard deviation of 398.3 cells /mm^3; the mean weight at baseline for HIV positive children was 19.13kg with a standard deviation of 9.5kgs, while the average baseline hemoglobin levels of children at baseline were 12.4 mm/Hg with a standard deviation of 4.15 mm/Hg. finally the average follow up time of children was 21.31months with a standard deviation of 18.75 months (see table2).

Table 2: Descriptive Statistics of Continuous Covariates at baseline for ART data set taken in Amhara region from 2010-2016

	BCD4 count	BWeight	Bhgb level	BAge	**Time**
Mean	465.1	19.13	12.4	6.5	21.31
Stand.dev	398.3	9.5	4.15	3.8	18.75

Where:

BCD4 count = baseline CD4 count

B Hgb level = baseline hemoglobin level

B Weight = baseline weight

B Age = baseline age

5.2 Exploratory data analysis

5.2.1 Individual profile plot for some selected patients

The individual profile plot for a sample of 100 subjects was done. From the transformed CD4 count, what can be noted from the panels is that generally there is evidence of between subjects variability as well as within subject variability. The subjects have largely variable CD4 values at the start and also possibly different evolutions over time; this suggests that perhaps linear mixed models with random intercepts and slopes could be plausible starting point. Also there is high within individual variability over time from both females and males (see figure3 and figure4).

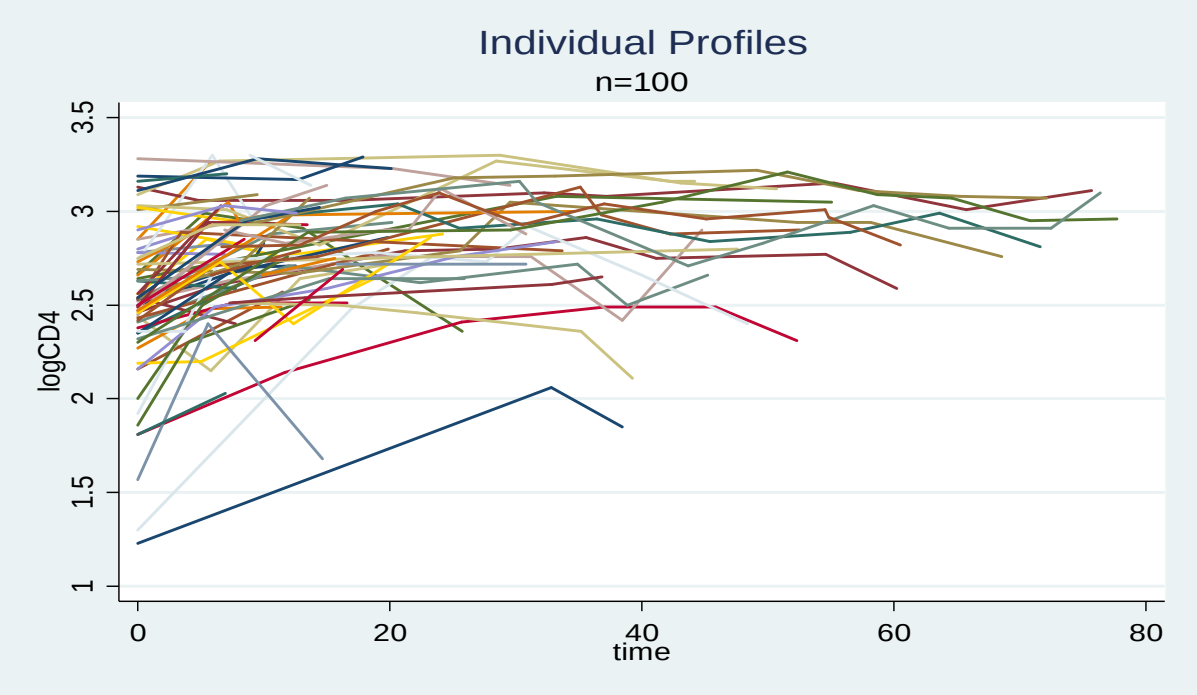


Figure 3: ***Individual profile plot for logCD4 count for ART data set taken from Amhara region from 2010-2016***

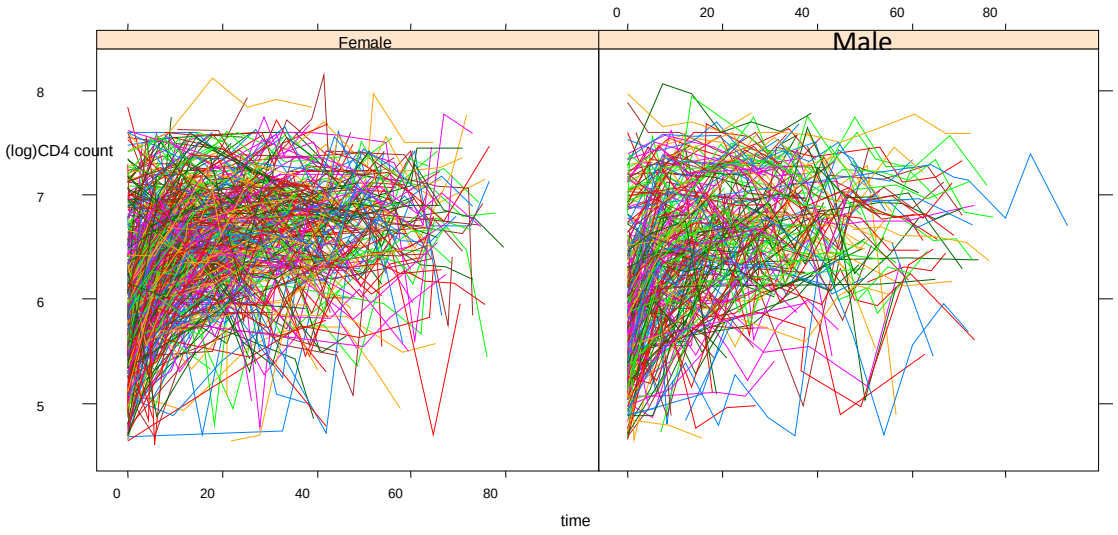


Figure 4: ***Individual profile plots for logCD4 count grouped by gender for ART data set taken from Amhara region from 2010-2016***

5.2. 2. Exploring the mean structure

The transformed CD4 (logCD4) values were used throughout in the exploratory data analysis. The data set was unbalanced in the sense that the number of repeated observations per individual was not the same for all patients and the measurements for all subjects were not taken at fixed time points of six monthly visits. The maximum numbers of measurements per subject were 12.

The average CD4 count of children at baseline were 465.1cells/mm^3 with a standard deviation of 398.28 cells/mm^3 , and it shows quick growth over time, further more the average rate CD4 count change of children is 5.01cells/mm^3 per month (see table3).

It is evident that the overall mean CD4+ count increases with time and stays stable over time. The mean CD4 count increment for males and females over time are similar (see figure5 and figure6).

The mean structure shown in (see appendix figure7), reveals that the average evolution of CD4 by ART adherence logCD4 counts seems to have a fast growth around 24 months of after ART initiation and continued stable over a certain months.

Table 3: Mean, standard deviation, and Number subjects in each time for ART data set taken from Amhara region from 2010-1016.

Time (month)	Mean	Std	N
0	465.08	398.28	951
1	713.50	439.00	769
2	800.93	484.916	618
3	886.89	604.63	478
4	877.07	453	379
5	932.25	602.69	297
6	994.33	1195	232
7	905.00	588.72	159
8	874.45	391.50	107
9	868.52	399.98	60
10	952.83	427.16	41
11	953.18	436.51	22

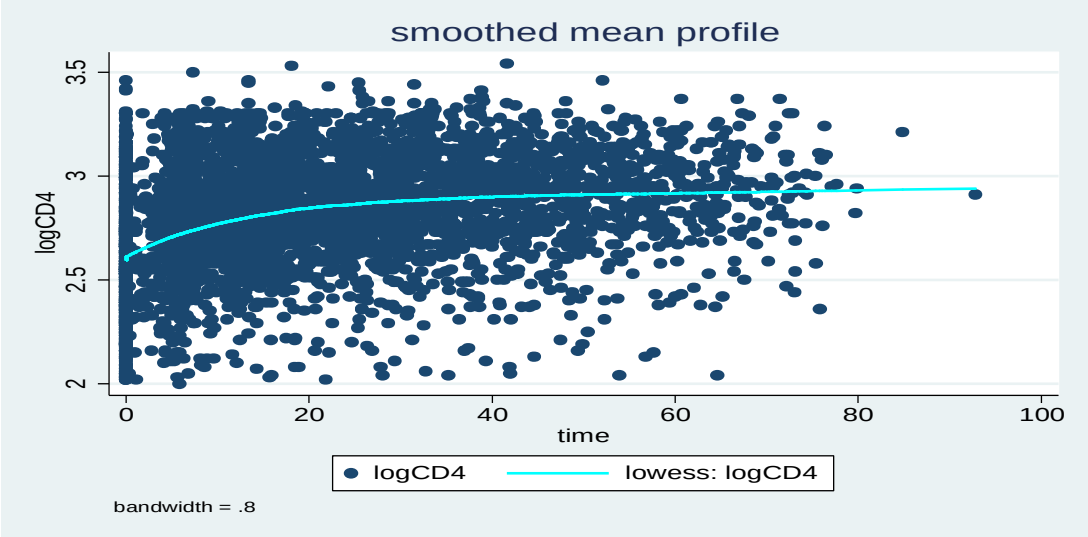


Figure 5: ***Smoothing Mean profile plots for the logCD4 count for ART data set taken from Amhara region from 2010-2016.***

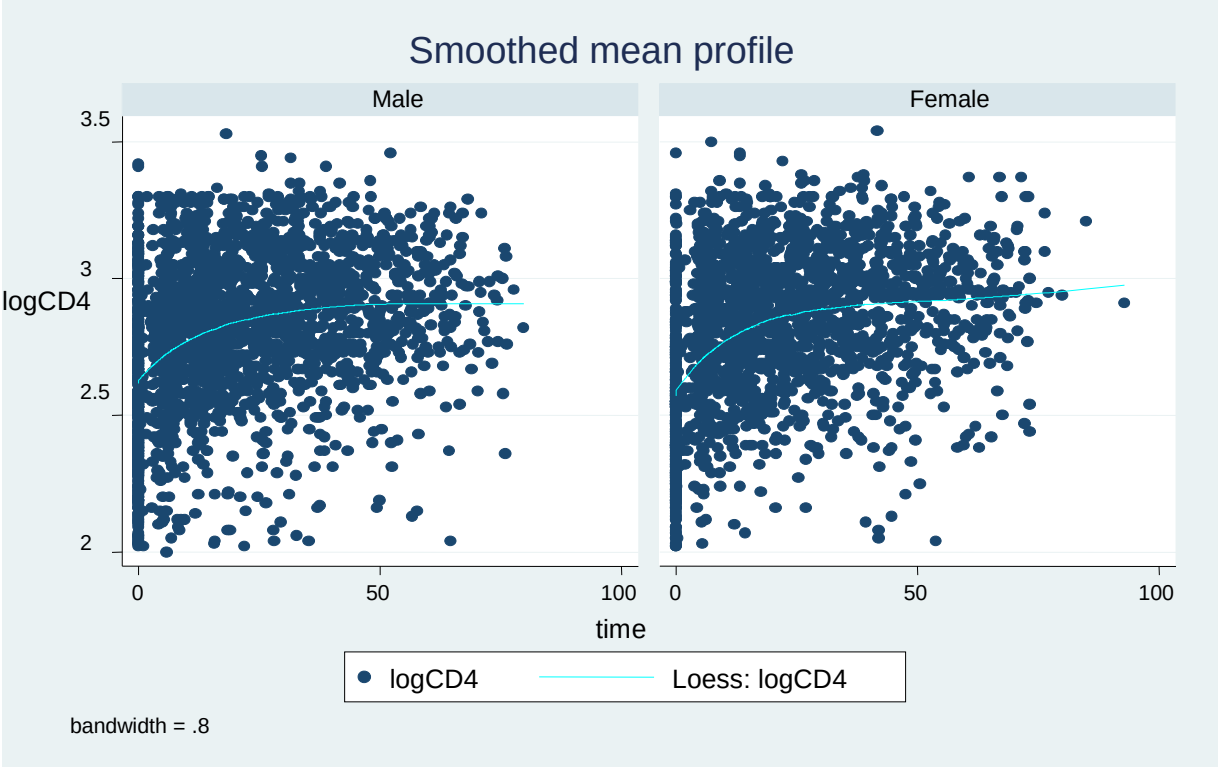


Figure 6: ***Smoothing mean profile grouped by gender for ART data set taken from Amhara region from 2010-2016***

5.2.3 Exploring correlation structure

The correlation matrix can be used to show the dependence between repeated measurements of the responses over time. The correlation matrix shown below revealed a positive correlation between any two repeated measurements.

Pair wise scatter plots were used for exploring the correlation between any two repeated measurements of CD4 count of patients and it appears that, there was a positive and linear relationship between CD4 counts at different time points (see figure7).

	CD40	CD41	CD42	CD43	CD44	CD45	CD46	CD47	CD48	CD49	CD410	CD411
CD40	1.0000											
CD41	0.6554	1.0000										
CD42	0.4640	0.7699	1.0000									
CD43	0.6377	0.8529	0.8791	1.0000								
CD44	0.6460	0.8725	0.5921	0.8795	1.0000							
CD45	0.7175	0.9697	0.8229	0.9359	0.9139	1.0000						
CD46	0.4778	0.8844	0.8489	0.9631	0.8898	0.9194	1.0000					
CD47	0.0386	0.7113	0.5962	0.7044	0.7374	0.6769	0.8652	1.0000				
CD48	-0.2674	0.5237	0.5292	0.4794	0.4609	0.4784	0.6661	0.8969	1.0000			
CD49	-0.0981	0.6336	0.5822	0.5758	0.590	0.612	0.728	0.887	0.9718	1.0000		
CD410	-0.1389	0.6432	0.502	0.507	0.570	0.568	0.703	0.924	0.973	0.968	1.0000	
CD411	0.0602	0.7492	0.7342	0.7021	0.6560	0.7349	0.8253	0.8928	0.9376	0.9775	0.9402	1.0000

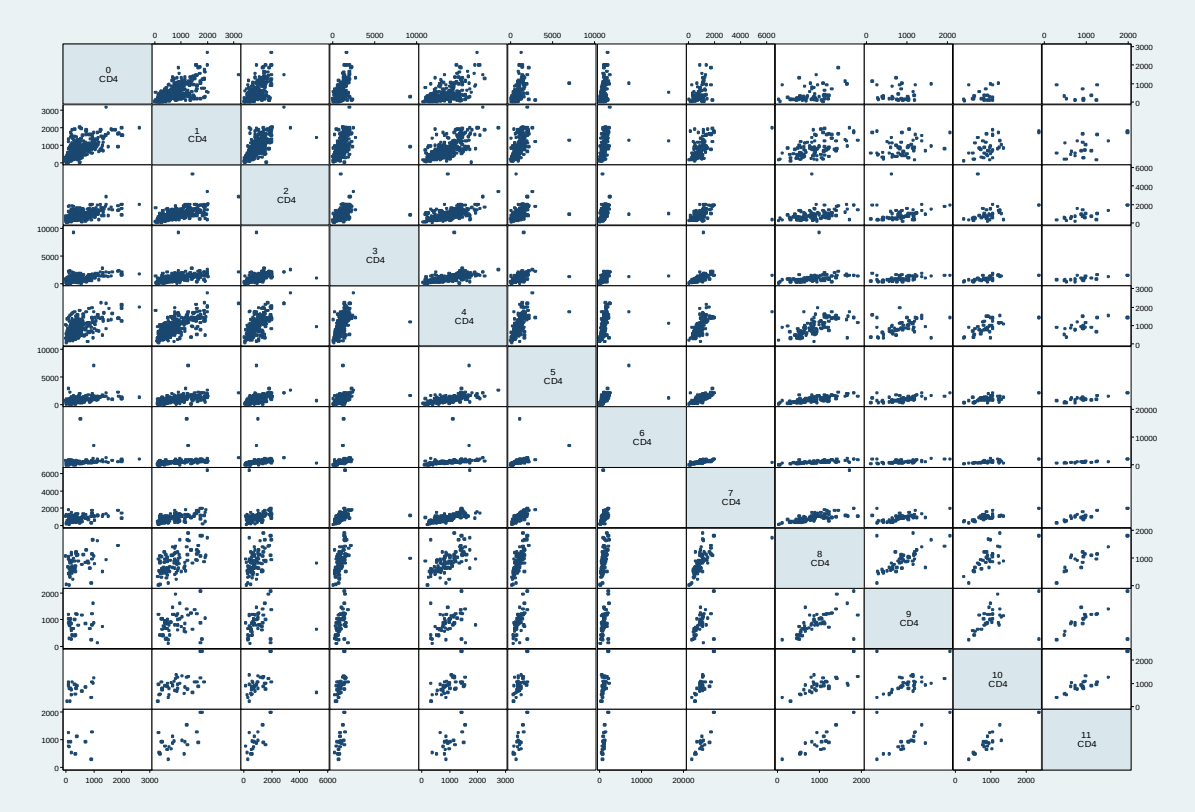


Figure 7: Scatter plot of correlation matrix for ART data set taken in Amhara region from 2010-2016

5.2 Linear Mixed Model

5.2.1 Univariate Linear Mixed Models

Covariates selected by manually baseline hemoglobin level, baseline CD4 count, ART adherence and gender is not significant effect on logCD4 count whose p-value > 0.2 from bivariate analysis, so we cannot enter in to multivariate analysis.

Table 4: *selection of correlation structure* for ART data set taken from Amhara region from 2010-2016

Correlation structure	AIC	BIC
Identity	12987.52	13099.52
Independent	12898.62	13021.88
Exchangeable	12982.36	13083.17
Un structured	12854.43	12955.23

Adding both random intercept and random slope are necessary to improve model fit which accounts extra between variability (see table 5).

Table 5: Random Effects Models with the associated values for the likelihood ratio test and p-value for ART data set taken from Amhara region from 2010-2016

Random effects	Likelihood-ratio test	**p-value**
Model1: intercepts	76.72	< 0.001
Model2: intercepts, time	35.48	< 0.001

Quadratic time function best express our model (see table6)

Table 6: tests for model extension for ART data set taken from Amhara region from 2010-2016

Model	REML log-likelihood
Linear	220.23
Quadratic	371.385
cubic	-33.08

5.2.2 Multivariate Linear Mixed Model

The results of multivariate analysis shows that variables baseline height, height, weight, and mothers ARV status were not significant predictor for the change in logCD4 count over time.

The average logCD4 count of children at baseline is 2.76 without consideration of covariate effects. Time have significant factor on logCD4 count $\beta 1 = 0.0058$ which expresses that the average logCD4 count of an individual increased by 0.0058 per month keeping the effect of other covariates constant. Baseline WHO clinical stages had undeniable effect on the logCD4 count change, that is, those patient who started ART at late clinical stages showed lower evolution of CD4 counts compared to those who started ART at earlier stages.

In addition the negative coefficients for all WHO stages refers to at base line, their mean CD4 counts are significantly lower than the reference group(i.e., $p < 0.001$ and < 0.001 for stage II,III respectively). However, over time all the three groups have significantly better average CD4 count as compared to the reference group.

For example, time by WHO stage II interaction ($\beta = 0.0076$) with entails the rate of increase in the logCD4 count for subjects in WHO stage II category is estimated to be 0.0076 times per month higher than the rate of increase among stage I patients.
Similarly, the rate of increase for subjects in stage III and IV categories are 0.0043 and 0.0059 times per month higher than the rate of increase in stage I patients respectively or the rates of change in the average logCD4 counts are increased by 0.0043 and 0.0059 times per month among stage III and IV patients correspondingly.

Furthermore, at base line the mean logCD4 count among children who have other opportunistic infection and disclosed to the disease and 0.088 times lower than the mean logCD4 count among children who have not opportunistic infection and do not disclosed to the disease (reference group) respectively.
Also age is strongly significant (p-value <0.001) risk factor for the change of logCD4 count which indicates that for every one year increase in age of children the average logCD4 count is decreased by 0.018 times.
In addition regimen type have undeniable effect on the change of logCD4 count, children take a combination of (d4T+3TC+EFV) drug lowers their logCD4 count by 0.098 times as compared to a children take (d4T+3TC+NVP) drug reference group keeping the effect of other covariates constant, while children take a combination of (AZT+3TC+EFV) drug gains logCD4 count by 0.02 times per month as compared the reference group (d4T+3TC+NVP) keeping the effect of other covariates constant.

Also interaction effect of ART adherence with time also significant effect on the logCD4 count change, which means that patients who have fair ART adherence over time gains an average logCD4 count 0.0059 times relative to poor ART adherence over time keeping the effect of other covariates constant. In addition patients who have good adherence over time gains an average logCD4 count 0.0057 per month relative to poor ART adherence keeping the effect of other covariates constant (see table7).

Table 7: ***multivariate analysis of linear mixed model for ART data set taken from Amhara region from 2010-2016***

Covariate	Estimate	p-value	95% CI
Constant	2.7579	0.000	2.580 2.936
Time	.0058	0.000	.004 .0076
$Time^2$	-.00018	0.000	-.000195 -.00016
WHO stage			
Stage I	0		
Stage II	-.0898	0.000	-.13367 -.045894
Stage III	-.1128	0.000	-.16048 -.06514
Stage IV	-.0492	0.366	-.15603 .05754
Hemoglobin level	.0102	0.018	.00175 .01864
Height	-.000394	0.738	-.00270 .00191
Weight	.00441	0.060	-.00019 .009008
Age	-.0181	0.000	-.02698 -.00926
Opportunistic infection			
No	0		
Yes	-.0476	0.034	-.09172 -.00354
Disclosure			
No	0		
Yes	-.0878	0.000	-.13497 -.04057
Baseline Height	-.0006895	0.600	-.00327 .00189
Baseline weight	-.00284	0.186	-.00705 .00137
Regimen type			
d4T+3TC+NVP	0		
d4T+3TC+EFV	-.09805	0.019	-.17969 -.01640
AZT+3TC+NVP	.02044	0.435	-.0308 .0717
AZT+3TC+EFV	.0713	0.030	.0069 .1356
TDF+3TC+EFV	.1704	0.037	.01040 .3305
Others	-.1056	0.567	-.4668 0.25567
WHO stage * time			
Stage II*time	.007590	0.000	.00587 .00930
Stage III*time	.00434	0.000	.0028 .00589
Stage IV*time	.00586	0.039	.000304 .0114
Regimen type*time			
d4T+3TC+EFV*time	.0015	0.0012	.0031 .00613
AZT+3TC+NVP* time	-.0036	0.001	-.0056 -.0015
AZT+3TC+EFV* time	-.0039	0.010	-.00689 -.00095
ART adherence			
Fair*time	.00597	0.004	.00500 .00693
Good*time	.00546	0.0023	.0051 .00634

The standard deviation of the intercept (b_{oi}= 0.276) this means that there is significant variation of the intercepts between study subjects which implies that evolution of logCD4 count vary across children at baseline.

It shows that the highest variability came from the random intercepts. It also shows that the standard deviation of the random intercepts was higher than that of the random slopes, pointing to higher between patient variability than the within patient variability. The total variability between individuals is estimated as $sd(b_o)$+ $sd(b_1)$ = 0.28 whereas the total variability within individual is 0.163.

However, the total variation in CD4 count is estimated to be 0.163+0.28= 0.443. The proportion of total variability that is attributed to within person variation is given by 0.163/0.443 is 36.7% while the proportion of total variability attributed to between individual variations in their general level of CD4 count is 0.28/0.443 is 63.21%. Therefore more than half of the variation is explained by the random intercepts and random slops (see table 8).

Table 8: Random parameter estimates for ART data set collected in Amhara region from2010-2016.

Random intercept parameters	Estimate	Std error	95% CI
$Sd(b_{oi})$	0.276	0.0076	.1892 .29812
$sd(b_{1i})$	0.004	0.00090	.00027 .00524
$corr(b_{1i}, b_{0i})$	0.729	0.2473	-.9611 .1070
$sd(\epsilon_{ij})$	0.163	0027607	.1578 .1687

5.3 Model diagnostic checking

In model diagnostic checking for longitudinal data analysis we use residual plots in order to evaluate the validity of model assumption. In normal linear model residuals are used to verify linearization, normality, and homoscedasticity of the errors and presence of outliers.

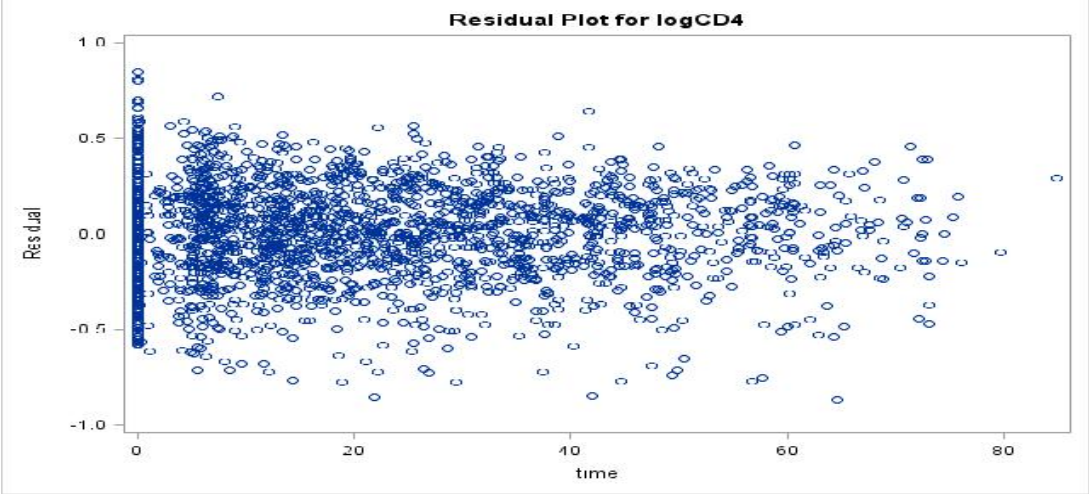
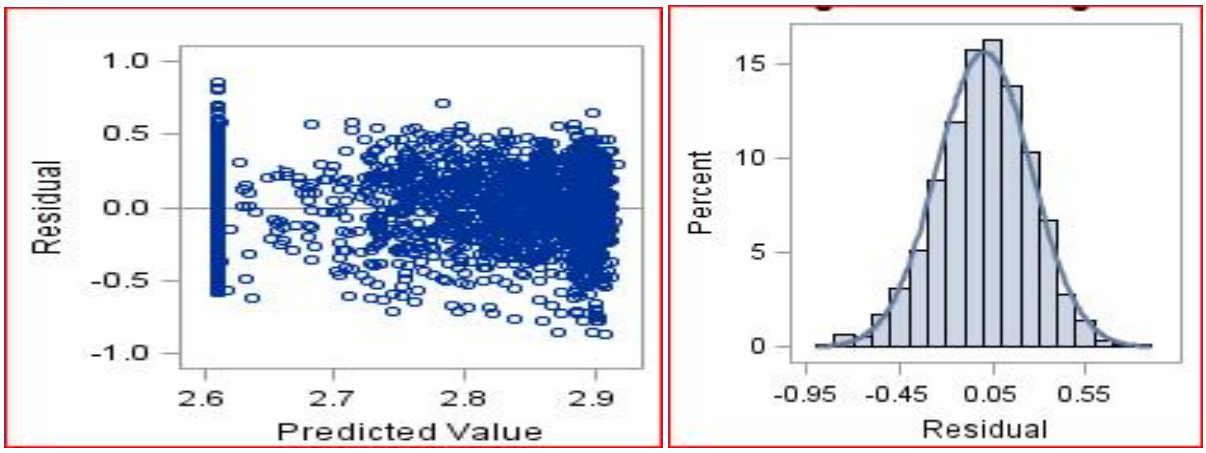
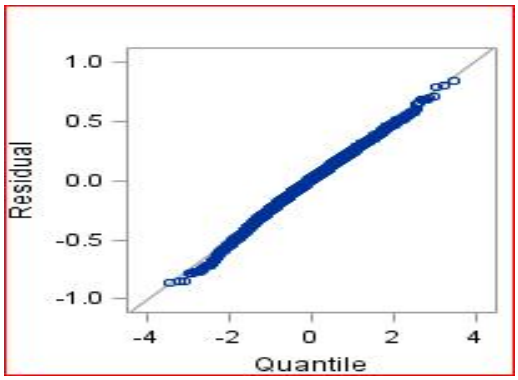


Figure 8: residual plot for logCD4 count for ART data set taken from Amhara region from 2010-2016

From this residual plot logCD4 count at each time point we understand that the observations bellow and above the fitted line are similar this indicates that the data is normally distributed with some sort of outliers.





The histogram of the residuals with overlaid normal density estimator and the normal quantile plot show that the residuals do not exhibit departure from normality. The “Residual fit” concentrated around zero implies that linear mixed model were well fitted the data.

6. Discussion

The aim of this study was to assess the CD4 count change over time among HIV positive children receiving antiretroviral therapy (ART) and its predictors. Since the data is correlated Linear Mixed Model (LMM) were applied for statistical modeling.

The ART data was analyzed using different plots (exploratory analysis) followed by model based outputs. From individuals profile plot, we observed the existence of variability in CD4 count within and between individuals. The exploratory analysis result for mean structure also suggested that on average, logCD4 tend to increase rapidly following the initiation of antiretroviral therapy consistent with a study[17] which is a reflection of the extent of suppression of viral replication, but it should be noted these plots are mean plots which can possibly be different from individual plots this may show some patients responding better than others.

The average rate change for CD4 count for this study was 30.04 six month or 60.08 per year consistent with a study [8, 27] who founds that an adequate CD4 response for most patients on therapy is defined as an increase in the range of 50–150 cells/mm3 per year with an accelerated response in the first 3 months of treatment

In addition the average logCD4 count increases in quadratic functions of time were found to be appropriate descriptions of the CD4 count change. This supports the results of [9, 28, 29] who found that after the patients initiated to the ART program their CD4 count increases and a dramatic decrease in viral load (VL) due to the therapy.

From multivariate linear mixed model we found covariates that have significant effect on the CD4 count change those covariates are opportunistic infection, disclosure, regimen type, WHO clinical stage, age, interaction effect of ART adherence, WHO clinical stage, and regimen type with time and time while baseline height, sex, baseline hemoglobin level, gender, mothers ARV status, have not significant effect on the change of CD4 cell count over time.

There was a significant negative effect of other opportunistic infections (β = -0.048; 95% CI -0.092, -0.0035), indicating that children who had other infections or opportunistic infections had lower levels of CD4 cell counts this is consistent with a study[8, 18] but

there was no significant difference between males and females in CD4 cell count change at baseline this supports figure3 as well as a study[8, 18, 30].

Regimen types have significant effect on the change of logCD4 count i.e. patients receiving two NRTI and one NNRTI shows low in logCD4 count gain at baseline this result directly consistent with a study done long term effect of HAART on the evolution of CD4 cell count among HIV positive children [11]. One of the possible reasons might be known hematological toxicity of the drug and combination of other drug suppresses CD4 count.

In addition disclosure has a negative effect on the change of CD4 cell count ($\beta = -0.088$, with p-value < 0.001) this may be due to frustration of children from disease this leads to stress and immunologic failure, but there is no studies related to association between CD4 cell count and disclosure in children.

On the other hand WHO clinical stage has negative significant effect at baseline i.e. Stage II and stage III but over time there is positive significant effect on the change of CD4 count over time this contradicts study conducted in Zimbabwe [30] on comparison of CD4 T-Cell changes in response to highly Active Antiviral Therapy (HAART) in adolescents and children enrolled at Parirenyatwa Hospital Family Care Centre which gets there is no significant effect on the change of CD4 cell count this may be due to comparison of study between children and adolescent for their study while, our study is studying the general trend of CD4 count change over time for all children, but this study is consistent with a study conducted in Ghana[31] who founds that the negative quadratic slope observed for children with more severe WHO clinical stages suggests a significant slowing in the increase in CD4 absolute counts for that age group.

In this study adjusted multivariate analyses, age at ART initiation remained negatively associated with logCD4 cell count ($\beta = -0.018$, 95% CI -0.027, -0.0093) this study is consistent with a study conducted in West and sub Saharan Africa [15, 17, 32]. This indicates that children at earlier age gains better CD4 recovery than older age. This implies that understanding the impact of treatment initiation at different CD4 counts and

age is essential to ensure that children reach adulthood with immune systems as unbroken as possible.

Hemoglobin level also associated with the improvement of CD4 count over time consistent with a previous study conducted in Gahanna[33] report shows that higher level of hemoglobin associated with an increased CD4 counts, improved immune reconstitution, and improved survival resulting in slower progression of disease; low level of hemoglobin may be associated with AIDS and death in children live with HIV.

Good and fair adherence marker of timelines of clinic attendance over time positively associated with the improvement of logCD4 count over time consistent with a study conducted in Bangladesh and India[9, 24] the explanation of this might be effective intake of ARV drug improves their CD4 count and enables to develop immunity and reduce virologic suppression.

The limitation of the study is that the causal association between CD4 cell count, viral load and other clinical parameters and over all treatment effect cannot be accurately determined due to retrospective nature of data which does not record all the necessary variables due to shortage of reagents or technical problems.

The strength of our study lies on the large number of children followed up along with long duration of follow up to six years.

7. Conclusion and recommendations

7.1 Conclusion

In conclusion the average rate of CD4 count change of children was originated on the expected level. The CD4 cell count change over time could be sufficiently described by a quadratic time function. The effects of several factors on the evolution were discussed and the influences of several covariates on the evolution of logCD4 cell count were indentified. Among these baseline WHO clinical stage, age, opportunistic infection, regimen type, Hemoglobin level and disclosure was the determinant factors that determine the CD4 count change in this study. In addition the interaction effect of ART adherence, regimen type, and WHO clinical stages with time are determinant predictors for CD4 count change over time.

7.2 Recommendations

The health care provider should made prompt management and diagnosis in order to prevent increased risk of opportunistic infection. The health center should ascertain counseling program to induce treatment adherence. Additional studies should conduct including other essential covariates that were not included in this study.

Reference

1. 20th international AIDS conference Melbourne, A., Global Fact Sheet: HIV/AIDS. 2014.
2. Control, C.f.D. and Prevention, HIV/AIDS basic statistics. Online at http://www. cdc. gov/hiv/basics/statistics. html. Accessed, 2014. **1**.
3. center for disease control and prevention, c.f.g.h., USAID, Strategies for Identifying and Linking HIV-Infected Infants, Children, and Adolescents to HIV Care and Treatment. 2017.
4. Vanisri H R1, V.N., CD4 Cell Count Analysis Among HIV Patients at a Tertiary Hospital in South India. Annals of Pathology and Laboratory Medicine, 2016. **03**(02).
5. Omobolawa Kukoyi1, L.R., Jonathan Powell3, Oliver Barry4, Meghan Prin5, Jonas Kusah2, Xiangyu Cong6 and Elijah Paintsil7, Viral load monitoring and antiretroviral treatment outcomes in a pediatric HIV cohort in Ghana. Kukoyi et al. BMC Infectious Diseases, 2016. **16**(58).
6. Amha Kebede (PhD), D.G.E.P.H.I., HIV Related Estimates and Projections for Ethiopia - 2014. Ethiopian Public Health Institute Federal Ministry of Health, 2014.
7. A.S. Foulkes, L.A., X. Li,M.A. Johnson,C. Smith,K. Mounzer,L.J. Montaner, Prediction based classification for longitudinal biomarkers. NIH Public Access, 2010.
8. Paintsil, E., et al., Absolute CD4+ T-Lymphocyte Count as a Surrogate Marker of Pediatric HIV Disease Progression. The Pediatric infectious disease journal, 2008. **27**(7): p. 629.
9. J. Prakash, B.D.M., Temporal changes of CD4 Count of HIV infected children using curve ffitting model. International Journal of Scientific and Research Publications, 2013. **3**(7).
10. M M Morsheimer, M.A.D., 2 MB ChB, MMed, Cert Paed ID; H Rabie,2 MB ChB, MMed, Cert Paed ID; M F Cotton,2 MD, PhD, Paediatric ART outcomes in a decentralised model of care in Cape Town, South Africa. SAJHIVMED, 2014. **15**(4).
11. Nash, D., et al., Long-term immunologic response to antiretroviral therapy in low-income countries: Collaborative analysis of prospective studies: The Antiretroviral Therapy in Lower Income Countries (ART-LINC) Collaboration of the International epidemiological Databases to Evaluate AIDS. AIDS (London, England), 2008. **22**(17): p. 2291.
12. Ginwalla R 2, C.E., Kawamya-Banda F1, Thomas R1, Mwiya M1, Kankasa C1, Prevalence of Clinical, Immunological and Virological Failure among Children on Haart at the University Teaching Hospital, Lusaka, Zambia. Medical Journal of Zambia, 2012. **39**(2).
13. Machado, D.M., et al., Factors associated with clinical, immunological and virological responses in protease-inhibitor-experienced Brazilian children receiving highly active antiretroviral therapy containing lopinavir-ritonavir. Brazilian Journal of Infectious Diseases, 2007. **11**(1): p. 16-19.

14. Vanisri, H. and N. Vadiraja, CD4 cell count analysis among HIV patients at a tertiary hospital in South India. Annals of Pathology and Laboratory Medicine, 2016. **3**(2): p. A81-86.
15. Sophie Desmondea, b., Fatoumata Dickoc, Fla Kouetad, Tanoh Ebouae, Eric Balestrea,b,, et al., Association between age at antiretroviral therapy initiation and 24-month immune response in HIV-infected children in West Africa. NIH Public Access, 2014. **28**(11).
16. Carolyn Bolton-Moore, M., et al., Clinical Outcomes and CD4 Cell Response in Children Receiving Antiretroviral Therapy at Primary Health Care Facilities in Zambia. JAMA, 2007. **298**(16).
17. Marie-Quitterie Picat1, 3., Joanna Lewis1,4*., Victor Musiime5, Andrew Prendergast6,7, Kusum Nathoo8,, P.N.N. Addy Kekitiinwa9, Diana M. Gibb6, Rodolphe Thiebaut2,3, A. Sarah Walker"6,, and R.C. Nigel Klein"1, 4, the ARROW Trial Team‘, Predicting Patterns of Long-Term CD4 Reconstitution in HIV-Infected Children Starting Antiretroviral Therapy in Sub-Saharan Africa: A Cohort-Based Modelling Study. PLoS Med, 2013. **10**(10).
18. Patel, K., et al., Long-term effects of highly active antiretroviral therapy on CD4+ cell evolution among children and adolescents infected with HIV: 5 years and counting. Clinical Infectious Diseases, 2008. **46**(11): p. 1751-1760.
19. Experimental Immunology Branch, H.a.A.M.B., 2 and Biostatistics and Data Management and N.C.I. Section, 3 National Institutes of Health, Bethesda, Maryland, Highly Active Antiretroviral Therapy in Human Immunodeficiency Virus Type 1-Infected Children: Analysis of Cellular Immune Responses. CLINICAL AND DIAGNOSTIC LABORATORY IMMUNOLOGY, 2001. **8**(5).
20. Jason A Mendoza, M., MPH1,2,*, Mogomotsi Matshaba, MB, BCh, BAO, BMedSci3,4, Jeremiah Makhanda, MSc3, Yan Liu, MS5, Matshwenyego Boitshwarelo, MD3, and Gabriel M. Anabwani, MBChB, MMed, MSCE, FRCPE3,4, Food Insecurity and CD4% among HIV+ Children in Gaborone, Botswana. NIH Public Access, 2014. **66**.
21. Edmonds, A.J., The effects of highly active antiretroviral therapy on survival and CD4 cell percentage in HIV-infected children in Kinshasa, Democratic Republic of Congo. 2011: THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL.
22. Murenjekwa, W., Comparison of CD4+ T-Cell changes in response to highly active antiviral therapy (HAART) in adolescents and children enrolled at Parirenyatwa Hospital Family Care Centre (2005-2010)-Secondary Data Analysis. 2013, University Of Zimbabwe.
23. al., M.D.e., Factors associated with clinical, Immunological and Virological Response imn Protease-Inhibitor- Experienced Brazillian Children Receiving HAART. The Brazillian Journal of infectious Diseases, 2007. **11**(1).
24. Mukherjee, A., et al., Outcome of highly active antiretroviral therapy in HIV-infected Indian children. BMC infectious diseases, 2014. **14**(1): p. 701.
25. Daniele de Brito Trindade(1), R.O., Leila D. Amorim(2),*, Choosing the right strategy to model longitudinal count data in Epidemiology: An application with CD4 cell counts. Epidemiology Biostatistics and Public Health, 2015. **12**(4).

26. Geert Molenberghs, G.V., A review on linear mixed models for longitudinal data, possibly subject to dropout. Statistical Modelling, 2001. **1**(235-269).
27. Hoffman, J., Role of the CD4 count in HIV management; CD4 count in resource limited settings. HIV Therapy, 2010. **4**(1).
28. Edmonds, A., et al., Quantification of CD4 responses to combined antiretroviral therapy over 5 years among HIV-infected children in Kinshasa, Democratic Republic of Congo. JAIDS Journal of Acquired Immune Deficiency Syndromes, 2012. **61**(1): p. 90-98.
29. Salvador Resino, J.M.B.n., 1 Dolores Gurbindo,2 Jose´ Toma´s Ramos,3 Juan Antonio Leo´n,5 M. Jose Mellado,4 and M. A´ ngeles Mun˜ oz-Ferna´ndez1, Viral Load and CD4+ T Lymphocyte Response to Highly Active Antiretroviral Therapy in Human Immunodeficiency Virus Type 1–Infected Children: An Observational Study. Clinical Infectious Diseases 2003. **37**(1).
30. Murenjekwa, W., Comparison of CD4+ T-Cell Changes In Response To Highly Active Antiviral Therapy (HAART) In Adolescents And Children Enrolled At Parirenyatwa Hospital Family Care Centre (2005-2010) - Secondary Data Analysis,. an MSc thesis 2013.
31. Barry, O., et al., Effectiveness of first-line antiretroviral therapy and correlates of longitudinal changes in CD4 and viral load among HIV-infected children in Ghana. BMC infectious diseases, 2013. **13**(1): p. 476.
32. Kovacs, A., et al., Immune reconstitution after receipt of highly active antiretroviral therapy in children with advanced or progressive HIV disease and complete or partial viral load response. Journal of Infectious Diseases, 2005. **192**(2): p. 296-302.
33. Louis Boafo Kwantwi, B.K.T., 1 Daniel Boateng,2 and Dan Yedu Quansah3, Body Mass Index, Haemoglobin, and Total Lymphocyte Count as a Surrogate for CD4 Count in Resource Limited Settings. Hindawi Journal of Biomarkers, 2017. **20**(17): p. 6.

Annexes I

A. Data abstraction form of HIV Treatment related to children from 2010 to 2016 follow up in Amhara Regional Hospitals, Ethiopia

Part I – Baseline Characteristics

S.No.	Characteristics / variables	Records
101	Facility/Hospital Name's	_____
102	MRN No.	_____
103	Date of Birth	___/_____/____
104	Age at enrollment in years (Month if <5 years)	_____
105	Sex	**1)** Male **2**) Female
106	Child's Address	Region _____Zone _____Woreda_____ Kebele _____
107	Primary care taker relationship	**1)** Parents **2)** Grandparents **3)** Relatives(sister/brothers…) **4)** Orphanage **5)** Others ___
108	Enrollment date (DD/MM/YYYY)	_____/_____/_____
109	HIV status Disclosure to child's	**1)** Yes **2**) No
110	Past OIs	**1)** Yes **2)** No if yes specify_____
111	Previous Child's mother ARV status	**1)** Yes **2**) No
112	If Q111 yes, which ARV drug or ART regimen	**1**) Sd-NVP **2)** AZT **3**) Combination ARV **4)** Other (specify) _____
113	ART start date (DD/MM/YYYY)	_____/_____/_____
114	First line ART regimen code	**1)** 4a **2)** 4b **3)** 4c **4)** 4d **5)** 4e **6)** 4f **7)** 4g
115	Base line Height (cm)	_____
116	Base line Weight (kg)	_____
117	Head circumference (cm) if child < 5	_____

	years	
118	Functional status of child	**1)** Working **2)** Ambulatory **3)** Bed ridden
119	Developmental status if child < 5 years	**1)** Appropriate for age **2**) Delay **3)** regression
120	WHO HIV clinical stages	**1)** I **2**) II **3)** III **4)** IV
121	Base line Hgb	_____

PART II: HIV care Follow up characteristics

	201	202	203	204	205	206	207		208	209	210	211
No.	**Date**	**Weight (kg)**	**Height/ HC**	**FS/ D.M**	**WHO stages**	**OIs**	**CD4/mm^3 or % if <5 Yrs**		**Hgb**	**ARV drugs**		
							Count	**%**		**Code**	**Side effects**	**Reason for change if**
0												
1												
2												
3												
4												
5												
6												
7												
8												
9												
10												
11												
12												
13												
14												
15												
16												
17												
18												
19												

	213	No. of visit based on Scheduled _____ No. of visit based Unscheduled _____ Total No. of visit _____										
No. of visit Unscheduled		1	2	3	4	5	6	7	8	9	10	11	12
No. of day difference per visit													
	214	Last status of patients **1.** Alive in first line **2.** Died **3.** Lost to follow up **4.** Drop **5.** Transfer to **6.** Switch to second line										

Data collector's: Name _____ signature _____ date_____
Supervisor's: Name _____signature _____date _____

Annexes II. Information Sheet and Consent Form

Information Sheet

Title of the research project: Predictors of the change of CD4+ cell counts among HIV infected children receiving antiretroviral therapy (ART) in Amhara region, Ethiopia. Retrospective longitudinal data analysis

Principal Investigator–Tilahun Yemanu (Bsc)

Advisors

1. Lemma Derseh (MPH, asst. professor)
2. Destaw Fetene (MPH)

Name of the organization: University of Gondar, College of Medicine and Health Sciences, Institute of Public Heath, Department of Epidemiology and Biostatistics

Sponsor: University of Gondar

Introduction

My name is Tilahun Yemanu I am post graduate student at University of Gondar in masters' degree in Biostatistics; I am doing a research to determine the change in CD4 cell count and predictors that affect the change in CD4 count among HIV infected children in Amhara region. A retrospective follow up study from HIV infected children in Amhara region, north Ethiopia is my study area. This research includes one principal investigator, ten data collectors, and three supervisors and two advisors from Gondar University.

Purpose of the Research

The aim of this study is to describe the change of CD4 cell count over time and its predictors among HIV infected children receiving antiretroviral therapy (ART) in Amhara region, Ethiopia.

Confidentiality and Anonymity

The information that I collect for this research will be kept confidential. Information about the patient that will be collected during the research will be stored in a file, which will not write their name or rather I use personal identification number.

Benefits: patients in this research project may not get direct benefit, but their benefit is likely to help us to assess the change of CD4 cell count over time among HIV infected children receiving antiretroviral therapy (ART) and factors that affect the change of CD4 cell count In addition finally, it will give an insight for policy makers and programmers to design new interventions.

Risks: in this study there is no risks occurred on the patient if we keep their confidentiality.

Persons to contact

This research will be reviewed and approved by the Ethical Review board of University of Gondar. If you wish to find about more or if you want to ask questions anytime you can use the contact addresses below:

1. Tilahun Yemanu :University of Gondar
   Tel: +251925978489 E-mail: yemanu.tilahun@gmail.com

2. Lemma Derseh (MPH, assist. professor)
   Tel: +251918773355
3. Destaw Fetene (MPH)
   Tel: +251918037193

Annexes III: Stata codes

############ convert data wide format in to long format###############

reshape long CD4 time WHO Weight Height Hgb FunctionalStatus Code Sideeffects daydiferrence , i(UniqueKey) j(t)

======\\\\\\\=========

summarize CD4

/**mean profile plot with loess smoothing/****

lowess logCD4 time, mean

lowess logCD4 Weight, mean

lowess logCD4 time, mean lineopts(lcolor(cyan) lwidth(medthick) lpattern(solid) connect(direct)) by(, title("smoothing mean profile by treatment adherance")) by(TA_2)

lowess logCD4 time, mean recast(line) lineopts(lcolor(purple)) by(, legend(on)) by(WHO)

lowess logCD4 time, mean lineopts(lcolor(purple)) by(, title("smmothing by WHO clinical stage")) by(, legend(on)) scheme(s2color) by(WHO)

lowess logCD4 time, mean lineopts(lcolor(purple)) by(, title("smmothing by ART regimen type")) by(, legend(on)) scheme(s2color)

/***individual profile plot /*****

xtline CD4 if UniqueKey <=100, overlay t(time) tlabel(#6) i(UniqueKey) legend(off) title("Individual Profiles") subtitle("n=100")

xtline CD4 if UniqueKey <=300, overlay t(time) tlabel(#6) i(UniqueKey) legend(off) title("Individual Profiles") subtitle("n=300")

xtline logCD4 if UniqueKey <=300, overlay t(time) tlabel(#6) i(UniqueKey) legend(off) title("Individual Profiles") subtitle("n=300"

xtline CD4 if WHOStage ==1 & UniqueKey <=250, overlay t(time) tlabel(#6) i(UniqueKey) legend(off) title("IndividualProfiles for WHO clinical stage1") name(stage1, replace)

xtline CD4 if WHOStage==2& UniqueKey <= 200, overlay t(time) tlabel(#6) i(UniqueKey) legend(off) title("Individual Profiles for WHO clinical stage2") name(stage2, replace)

xtline CD4 if WHOStage==3& UniqueKey <= 200, overlay t(time) tlabel(#6) i(UniqueKey) legend(off) title("Individual Profiles for WHO clinical stage3") name(stage3, replace)

xtline CD4 if WHOStage==4& UniqueKey <= 1000, overlay t(time) tlabel(#6) i(UniqueKey) legend(off) title("Individual Profiles for WHO clinical stage4") name(stage4, replace)

xtline CD4 if FunctionalStatus ==1& UniqueKey <= 1000, overlay t(time) tlabel(#6) i(UniqueKey) legend(off) title("Individual Profiles for FunctionalStatus") name(bedriden, replace)

/*** linear mixed model /****

xtmixed logCD4 time c.time#c.time c.time#c.time#c.time Weight Height i.FunctionalStatus i.OI i.Primarycaretaker i.Disclosure i.MARVstatus BHeight Bweight i.WHOStage i.sex#c.time i.Code#c.time i.TA_2#c.time i.WHO#c.time i.OI#c.time i.Disclosure#c.time || UniqueKey:time Height, cov(un) reml nolog

estat ic

for multiple imputation####

mi set mlong

mi register imputed Weight

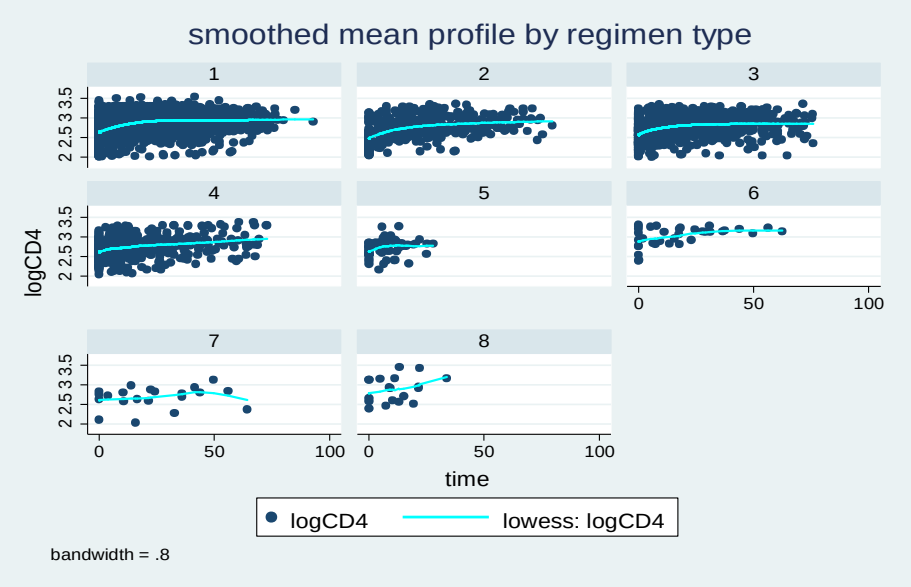
mi impute regress Weight, add(5) rseed(1234)

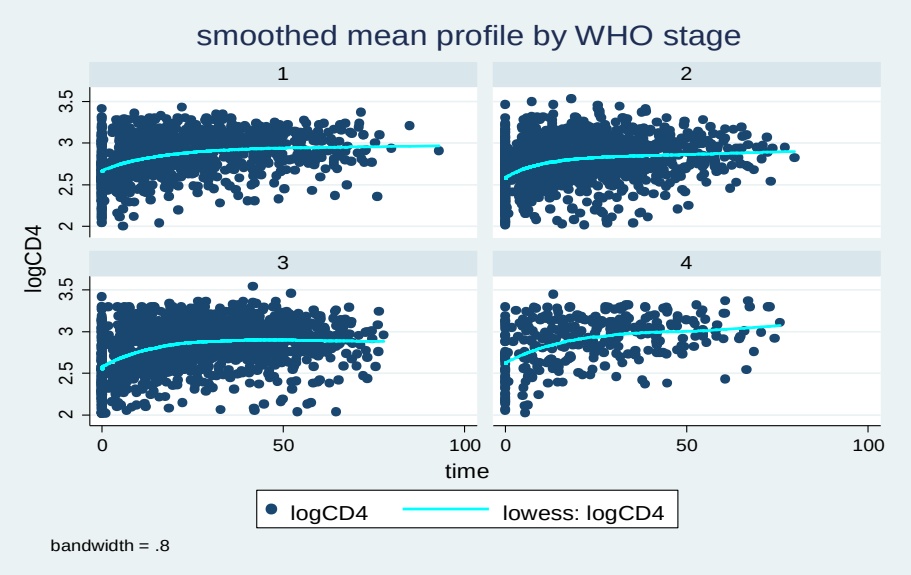
mi estimate: xtmixed logCD4 time Age i.sex Weight Height Hgb i.OI i.Disclosure i.MARVstatus || UniqueKey: time, cov(un) reml nolog

xtmixed logCD4 time Age Hgb i.OI#c.time i.Baselinecode i.Code#c.time i.Disclosure#c.time i.MARVstatus i.TA_2#c.time i.WHO#c.time ||UniqueKey: time Age Hgb, cov(un) reml nolog

xtmixed logCD4 Age time i.OI i.Disclosure i.WHOStage i.WHO#c.time Hgb Height Weight c.Height#c.time i.FunctionalStatus#c.time i.Baselinecode#c.time ||UniqueKey: time Hgb Age, cov(un) reml nolog

Annexes IV: some selected stata outputs





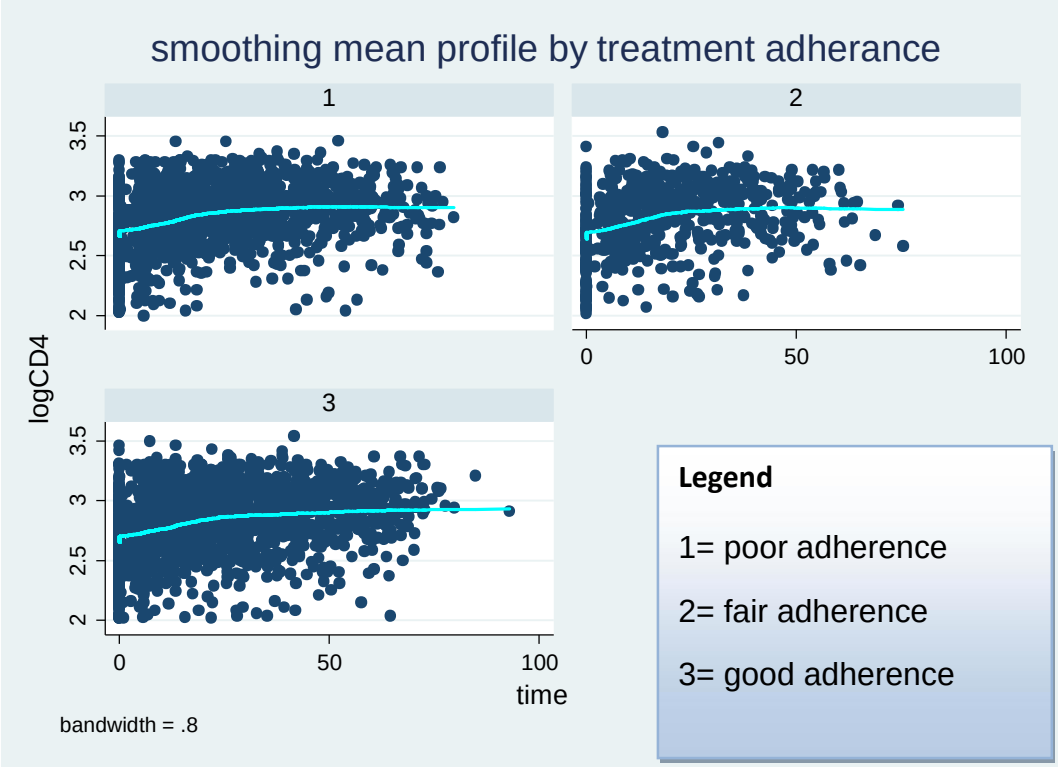


Figure 9: ***Smoothing mean profile by ART adherence for ART data set taken from Amhara region from 2010-2016***

Table 6: univariate analysis for ART data set taken from Amhara region from 2010-2016

Covariates |	Coef (Std. Err).	z	P>|z|	[95% Conf. Interval]
cons |	2.703(.0115)	35.32	0.000	2.680 2.7254
time |	.119(.005)	24.45	0.000	.1101806 .129384
(time)2 |	.0000488(3.05e-6)	15.97	0.000	.0000428 .0000548
sex(female) |	-.0042(.0149)	-0.28	0.780	-.0334 .0251
Age |	-.0289(.0018)	-16.00	0.000	-.0324892 -.0253983
Height |	-.0027(.0003)	-8.27	0.000	-.0032855 -.0020263
BHeight |	-.00068(.00019)	-3.59	0.000	-.0010589 -.0003112
Bweight |	-.0079 (.00079)	-10.02	0.000	-.009464 -.006368
Hgb |	.00142(.0013)	1.10	0.272	-.0011111 .003940
BaselinHgb |	-.00109(.0012)	-0.93	0.353	-.0033912 .001211
WHOStage (stage1 refe)				

stage2 |	-.0492(.0188)	-2.61	0.009	-.0862	-.0123
stage3 |	-.0415(.0191)	-2.18	0.030	-.0789	-.00414
stage4 |	.0205 (.0309)	0.66	0.507	-.0401	.08114
Bregimentype |					
2 |	-.0975(.0274)	-3.56	0.000	-.1511	-.04376
3 |	-.0503(.017)	-2.92	0.004	-.0839	-.01650
4 |	-.074(.0225)	-3.29	0.001	-.1178	-.0298
5 |	-.0357(.0513)	-0.70	0.487	-.1361	.06485
6 |	.180(.0726)	2.48	0.013	.0378	.3224
OI (yes) |	-.058(.0169	-3.43	0.001	-.0913	-.02497
MARVstatus(yes) |	.1092(.0376)	2.91	0.004	.0356	.1834
Disclosure(yes) |	-.127(.0177)	-7.21	0.000	-.16195	-.0927
Primarycaretaker(parents refe)					
GParents |	-.0635(.0287)	-2.22	0.027	-.1197	-.0074
Relatives |	-.0311(.0342)	-0.91	0.363	-.0983	.03599
Orphanage |	-.1442(.0470)	-3.07	0.002	-.2363	-.0520
FunctionalStatus(working refe)					
Ambulatory |	.0667(.0163)	4.09	0.000	.0348	.09864
Bedridden |	. 0147(.106)	0.14	0.890	-.1928	.2222

Declaration

I, the undersigned, senior MPH student declare that this thesis report is my original work in partial fulfillment of the requirement for the degree of Master of public health in Biostatistics.

Name: Tilahun Yemanu

Signature: _____

Place of submission: institute of Public Health, college of Medicine and Health sciences, university of Gondar.

Date of submission: _____

ASSURANCE OF INVESTIGATOR

The undersigned agrees to accept responsibility for the scientific, ethical and technical conduct of this research project and for provision of required progress reports as pre terms and conditions of the research and publications office of the University of Gondar.

Name of the student: Tilahun Yemanu_____

Date: 05/06/2017_____ Signature: _____

Approval of the advisor (s)

Advisors:

Name	Signature	Date
1. _____	_____	_____
2. _____	_____	_____